



# De l'image vers la compression

Charles Wagner

## ► To cite this version:

Charles Wagner. De l'image vers la compression. [Rapport de recherche] RR-2035, INRIA. 1993. inria-00074636

**HAL Id: inria-00074636**

**<https://inria.hal.science/inria-00074636>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***De l'image vers la compression***

Charles Wagner

**N° 2035**

Septembre 1993

## PROGRAMME 1

Architectures parallèles,  
bases de données,  
réseaux et systèmes distribués

 ***apport  
de recherche*****1993**





## De l'image vers la compression

Charles Wagner\*

Programme 1 — Architectures parallèles, bases de données, réseaux  
et systèmes distribués  
Projet API

Rapport de recherche n°2035 — Septembre 1993 — 38 pages

**Résumé :** Les domaines d'applications mettant en jeu l'image vidéo ne cessent de s'étendre sous l'impulsion des progrès réalisés en traitement du signal, en architecture de machines ainsi que des avancées technologiques en matière d'intégration de composants. En télévision haute définition cette évolution est plus particulièrement sensible et l'on constate que l'application, les algorithmes mis en œuvre, les supports de transmission utilisés et les aspects normalisation sont étroitement liés.

Dans le contexte de l'étude d'une machine de traitement vidéo programmable il est apparu nécessaire de posséder une vision globale de ces différents aspects. Ce rapport concrétise cette étude et se veut avant tout être un état des lieux, à un instant donné, du domaine d'application cité.

**Mots-clé :** télévision, codage

*(Abstract: pto)*

\*wagner@irisa.fr

Unité de recherche INRIA Rennes  
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)  
Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 38 38 32

# Going from Video Signals to Compression

**Abstract:** The application domains for video imaging are ceaselessly expanding under the combined impulses of progress in signal processing and video architectures, as well as technological advances in the integration of components. In high-definition television this evolution is particularly noticeable, and one notices that the applications, the algorithms implemented, the transmission equipment used and the standardization of formats are closely coupled.

In the context of a study of a programmable video coder it has become necessary to have a global vision of all these aspects. This report formalizes this study and presents above all a state-of-the-art, at a given moment, of the domain of the given application.

**Key-words:** television, coding

## Introduction

La vidéo digitalisée et codée est parmi nous. Elle est en train de révolutionner un grand nombre d'applications telles que disque laser digital, caméra électronique, système de vidéo-téléphone et de vidéo-conférence, outils vidéo interactifs sur PC et stations de travail, programmes délivrés sur le câble et le satellite.

Toutes ces applications, à des degrés divers, nécessitent de réaliser l'adéquation, par l'utilisation de techniques de compression spécifiques, entre deux contraintes antagonistes à savoir la largeur de la bande passante du support de transmission et la qualité de l'image transmise.

Pour maîtriser cette explosion, des standards concernant les différents aspects de la transmission et de la réception de la vidéo se sont développés sous l'égide des organismes internationaux et sont l'aboutissement de longues années de recherche, d'évaluation et de compétition entre les grands groupes industriels. Dans la mesure du possible ces standards permettent que cette révolution puissent se faire dans la continuité sans remettre en cause l'existant.

La première partie de cet article est consacrée à la production d'images conventionnelles à la fois sous les aspects analyse vidéo de l'image et présentation des standards de transmission actuels dont les limitations peuvent être réduites par le système MAC (*Multiplex Analog Components*). La télévision haute définition ainsi que la numérisation du signal vidéo sont abordées dans la deuxième partie. On décrit ensuite, dans la troisième partie, les différents supports de transmission de l'image tant sous l'angle des débits que des types d'applications visées. Après une présentation de différents projets internationaux relatifs au codage de la vidéo, on passe en revue les principaux algorithmes normalisés utilisés en codage d'images animées.

## 1 Production d'images TV

### 1.1 Généralités

#### 1.1.1 Analyse de l'image

En télévision [3] l'image est analysée séquentiellement selon un balayage horizontal de gauche à droite. Le nombre de lignes détermine la définition.

La manière, dont l'image est analysée, est influencée par certaines caractéristiques physiologiques de l'oeil humain. Pour éviter le phénomène de scintillation, la

fréquence d’affichage des images va être reliée au phénomène de persistance rétinienne, de l’ordre du 1/15ème de seconde, qui fait que l’oeil ne peut séparer deux stimulations visuelles trop rapprochées. Par ailleurs la définition de l’image sera bornée par le pouvoir séparateur limité de l’oeil.

Aux considérations précédentes s’ajoutent des contraintes techniques telles que les interférences entre la fréquence trame et la fréquence du secteur (50Hz ou 60Hz selon les pays) et les facilités de production par division des fréquences ligne.

On transmet donc 25 images par seconde pour reconstituer le mouvement. Pour éviter le scintillement qui en résulte, on utilise la technique du balayage en deux trames entrelacées, lignes impaires puis lignes paires soit 50 trames ou demi-images par seconde.

Le balayage progressif réalise une analyse ligne par ligne. Il regagne actuellement de l’intérêt en raison des traitements numériques de l’image qu’il facilite. En effet le balayage entrelacé mélange les informations spatiales et temporelles alors que le balayage progressif sépare complètement les natures spatiales et temporelles de l’information ce qui conduit à une détection et une estimation de mouvement performante.

Selon les compromis effectués on va aboutir à des systèmes de télévision ayant des fréquences trame, ligne ainsi que des définitions différentes.

Le signal vidéo, pour des images noir et blanc, est donc constitué d’un signal de luminance correspondant à l’intensité lumineuse de l’image analysée auquel sont associés les signaux de synchronisation et de suppression de lignes et de trames.

### 1.1.2 Couleur

Une couleur est caractérisée par une information de luminance (intensité lumineuse) et des informations de chrominance telles que la teinte (fréquence couleur) et la saturation (degré de pureté ou largeur de bande de la fréquence couleur). Elle peut être définie par synthèse des trois couleurs fondamentales rouge (R), vert (V) et bleu (B). Le signal vidéo couleur sera composé d’une information de luminance Y qui est une combinaison linéaire des informations R, V, B ainsi que de deux signaux de chrominance  $DR = R - Y$  et  $DB = B - Y$ . L’intérêt d’utiliser des signaux de différence provient du fait que ceux ci deviennent nuls chaque fois que l’image analysée est blanche, noir ou grise c’est à dire  $R = V = B = k$  avec  $0 \leq k \leq 1$  (se reporter à la notion de compatibilité).

### 1.1.3 Standards mondiaux

Les fréquences réservées à la télévision ont été fixées par les Conférences Européennes de Radiodiffusion : bande I (41 à 68 Mhz), bande III (162 à 216 Mhz), bande IV et V (470 à 960 Mhz). A l'intérieur de ces bandes existe un découpage en canaux correspondant à la portion de fréquences attribuée à un émetteur. La portée normale des ondes hertziennes en V.H.F. et U.H.F. est relativement réduite (100km) d'où la nécessité d'un grand nombre de réémetteurs qui peuvent donc émettre sur le même canal s'ils sont suffisamment éloignés les uns des autres.

Les différents standards mondiaux sont définis par les critères suivants : fréquence ligne, fréquence image, largeur du canal, bande passante vidéo, type de modulation, écart entre porteuse son et vidéo. Ces standards sont repérés par les lettres A à L (standard français).

### 1.1.4 Compatibilité

En télévision noir et blanc l'image est analysée séquentiellement par points en transmettant un signal électrique proportionnel à la luminance du point. On ajoute les signaux de synchronisation permettant l'alignement des points en lignes et des lignes en trames. En télévision couleur il faut ajouter, à la luminance, la chrominance (couleur et saturation) de chaque point. Par ailleurs le système de transmission utilisé doit assurer la compatibilité directe, recevoir en noir et blanc une émission couleur sur un poste noir et blanc, et la rétrocompatibilité, recevoir en noir et blanc une émission noir et blanc sur un poste couleur.

La scène est analysée par trois tubes comportant chacun un filtre pour les 3 couleurs fondamentales (R, V, B). Chacune des tensions résultantes (ER, EV, EB) est affectée d'un coefficient correcteur **gamma** pour tenir compte des caractéristiques tension-lumière du tube. Le signal de luminance transmis  $Y = 0,3 \cdot ER + 0,59 \cdot EV + 0,11 \cdot EB$  tient compte de la courbe de sensibilité de l'oeil. Ce signal module l'émetteur H.F. avec la bande passante normale (celle des émetteurs noir et blanc définie précédemment) et l'on assure ainsi la compatibilité. Il reste donc à transmettre les signaux de chrominance associés à la luminance et ceci dans une bande passante identique (6,5 Mhz) à celle des émetteurs de télévision noir et blanc, sans dégradation notable de la qualité visuelle de l'image.

L'insertion des signaux de chrominance dans la bande passante du signal de luminance, avec un minimum d'interférence, a été rendu possible par les carac-



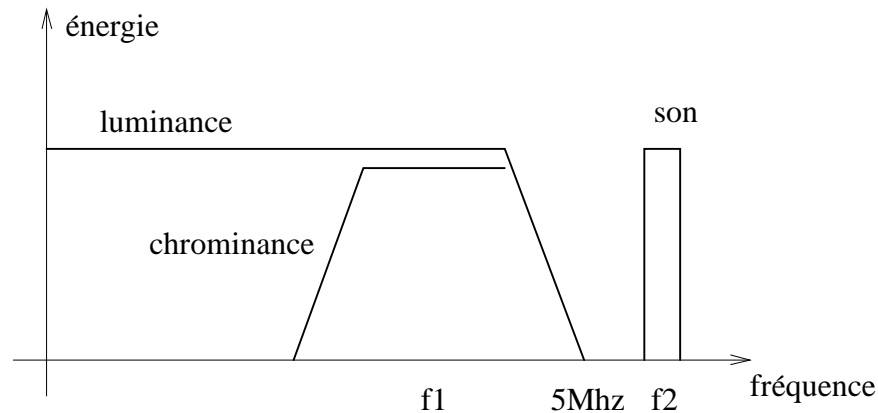


Figure 1 : spectre de fréquence du signal TV couleur

téristiques suivantes. L'oeil a un faible pouvoir séparateur en vision colorée et la finesse de l'image est obtenue grâce à la large bande passante du signal vidéo. La bande passante utilisée pour les signaux de chrominance peut être réduite à 3 Mhz. Par ailleurs l'énergie du signal de luminance est concentrée dans les basses fréquences. Enfin les fréquences du spectre de luminance sont des multiples de la fréquence de balayage ligne laissant les zones intermédiaires libres. La solution adoptée consiste à moduler par les signaux de chrominance une sous-porteuse à bande étroite, de fréquence un multiple entier et impair de la demi-fréquence de balayage horizontal.

C'est sur ce principe (figure 1) que sont basés les différents systèmes de codage de la couleur tels que PAL (Phase Alternating Line), SECAM (Séquentiel Couleur A Mémoire), NTSC (National Television System Committee). Ils diffèrent essentiellement par la manière dont sont modulées les deux informations de chrominance.

## 1.2 Principaux standards

- SECAM (figure 2) format image 4:3, 50Hz, 625 lignes, bande passante vidéo 6Mhz, largeur de canal de transmission 8Mhz, les deux signaux de chrominance sont transmis l'un après l'autre, une ligne sur deux d'où le nom de séquentiel.

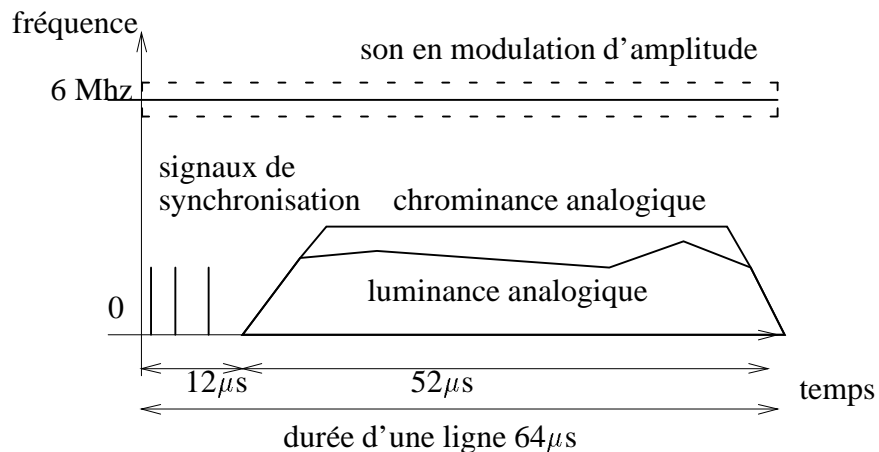


Figure 2 : signal SECAM

- NTSC : format image 4:3, 60Hz, 525 lignes, bande passante vidéo=4Mhz, largeur de canal 6Mhz ; pour la sous-porteuse de chrominance la phase exprime la teinte et l'amplitude la saturation ; en noir et blanc il n'y plus de sous-porteuse.
- PAL : format image 4:3, 50Hz, 625 lignes, bande passante vidéo 5Mhz, largeur de canal 8Mhz ; une porteuse pour le son modulée en amplitude et une porteuse pour l'image avec une tranche de  $12\mu s$  pour la synchronisation et une tranche de  $52\mu s$  pour le signal de chrominance et de luminance multiplexés à l'émission (modulation de fréquence analogique), même mode de transmission de la chrominance qu'en NTSC mais inversion de la phase une ligne sur deux.

En SECAM L la durée du balayage ligne est de  $52,5\mu s$  à 25 images/s pour 625 lignes soit  $64\mu$  au total par ligne. En réalité l'image utile est de 575 lignes, la durée correspondant aux 50 lignes restantes étant utilisée pour les signaux de suppression de trame, d'égalisation et de synchronisation.

### 1.3 Limitations des systèmes actuels

Une première limitation est celle imposée à la finesse de l'image. Le nombre de lignes est insuffisant pour les grands écrans actuellement commercialisés. La bande passante vidéo nécessaire pour transmettre une résolution horizontale correspondant à la résolution verticale varie comme le carré du nombre de lignes. Le mode de transmission de la couleur, par l'emploi de sous-porteuses plus ou moins mélangées avec la partie visible de l'image a comme conséquence une pollution de l'image : pointillés, structures parasites se déplaçant sur les contours, transitions verticales de couleur, entre plages de couleurs, médiocres. La structure en trames entrelacées cause un phénomène de scintillement sur les contours quasi horizontaux de l'image. Par ailleurs le son est de mauvaise qualité (modulation d'amplitude) et le système de transmission est limité à un seul signal son.

La résolution et la taille de l'image ne peuvent être séparées. L'expérience montre que plus l'image est grande et plus un format large devient attrayant (format 16/9). A l'origine la télévision a été mise sur pied en privilégiant de façon quasi-exclusive la diffusion hertzienne sur émetteurs terrestres. Les systèmes MAC/paquets permettent, au moins sur satellites et câbles, de s'affranchir de certains défauts du SECAM. Il est possible par des systèmes de modulation transportant des informations numériques d'améliorer de façon appréciable l'efficacité d'utilisation des bandes de fréquences affectées à la télévision terrestre. Une telle approche par diffusion numérique constitue une remise en cause de la gestion des fréquences et des récepteurs.

### 1.4 Famille MAC

Lorsqu'il est apparu possible de diffuser des programmes de télévision par satellite il est devenu évident qu'il fallait définir de nouvelles normes prenant en compte un certain nombre d'impératifs :

- meilleur utilisation de la puissance du satellite ;
- améliorer la qualité de l'image par rapport au SECAM ou PAL ;
- augmenter le nombre et la qualité des canaux sonores (stéréophonie, multilinguisme) ;
- permettre une gestion des ressources radio-fréquences assurant des extensions vers un format plus proche de celui du cinéma ;

- réutiliser au maximum le parc de téléviseurs existant ;
- adopter un décodeur unique capable de recevoir des signaux reçus par satellite ou faisceau hertzien ;

### 1.4.1 Historique

En 1981 les ingénieurs de l'IBA (*Independent Broadcasting Authority*) ont mis au point le système de transmission MAC. Ils ont proposé de transmettre successivement l'information de chrominance puis de luminance en les comprimant de manière à ce que ces deux informations soient transmises dans la durée normale de balayage d'une ligne (figure 3). Ces informations sont transmises sous forme d'un signal analogique. On obtient un multiplexage temporel contrairement au SECAM ou au PAL qui comporte un multiplexage en fréquence. L'intervalle de temps correspondant au retour de ligne et de trame est utilisé pour transmettre des informations numériques de son et de service. Il est possible d'associer plusieurs voies sonores à chaque image pour la diffusion d'un commentaire en plusieurs langues. Les informations sont transmises en paquets, chacun muni d'une étiquette selon sa nature d'où le nom du système MAC paquets.

### 1.4.2 Différents systèmes MAC

Plusieurs systèmes MAC ont été proposés :

- B-MAC américain utilisé par l'Australie, simple mais n'ayant que trois canaux son
- C-Mac ayant la préférence de l'U.E.R., système anglais offrant un meilleur service (8 canaux son monophonie) nécessitant une mise en oeuvre coûteuse (démodulation son et image distincte). Le B-MAC et le C-MAC sont incompatibles avec les réseaux câblés à bande étroite (7-8 Mhz).
- D-MAC ayant la préférence des industriels européens (Philips et Thomson) et qui a été retenu par la France et l'Allemagne.

### 1.4.3 D2-MAC

Cette norme, correspondant à un D-MAC modifié, a été adoptée par la France et l'Allemagne pour les satellites TDF1, TDF2 et TV-SAT. Il existe en fait deux normes D2-MAC, une pour la transmission hertzienne et le câble utilisant la modulation en amplitude et la seconde pour le satellite utilisant la modulation de fréquence, moins sensible au bruit mais plus coûteuse en bande passante. La bande passante requise pour une transmission D2-MAC est de 8,4 Mhz.

**Multiplexage temporel** Les informations de luminance (Y) et de chrominance (U, V), alternativement U pour une ligne puis V pour la suivante, sont compressées pour tenir sur les  $52 \mu s$  correspondant à la durée du trajet aller du spot sur une ligne. Ces informations sont transmises en mode analogique. Par contre les opérations de compression (à l'émission) et de décodage-décompression (à la réception) sont effectuées sur le signal numérisé. Le signal est échantillonné, éventuellement mémorisé sous forme numérique (chrominance séquentielle), traité (filtrages verticaux), (dé)comprimé et converti en analogique pour la transmission ou la visualisation. Le taux de compression de la luminance est de 3:2 (5:4) et celui de chrominance de 3:1 (5:1).

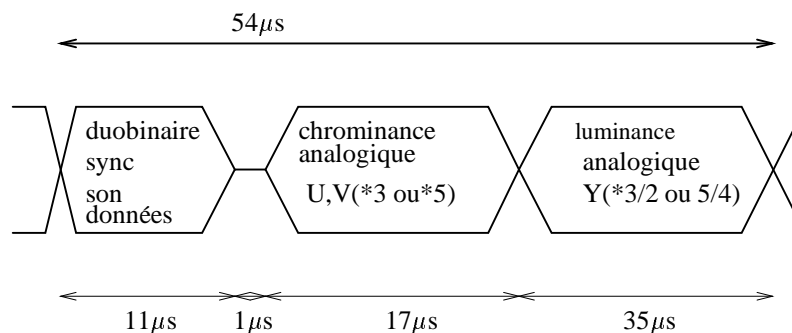


Figure 3 : structure d'une ligne MAC

Les  $12 \mu s$  correspondant au retour de ligne sont utilisées pour la transmission en numérique du son et des informations de service.

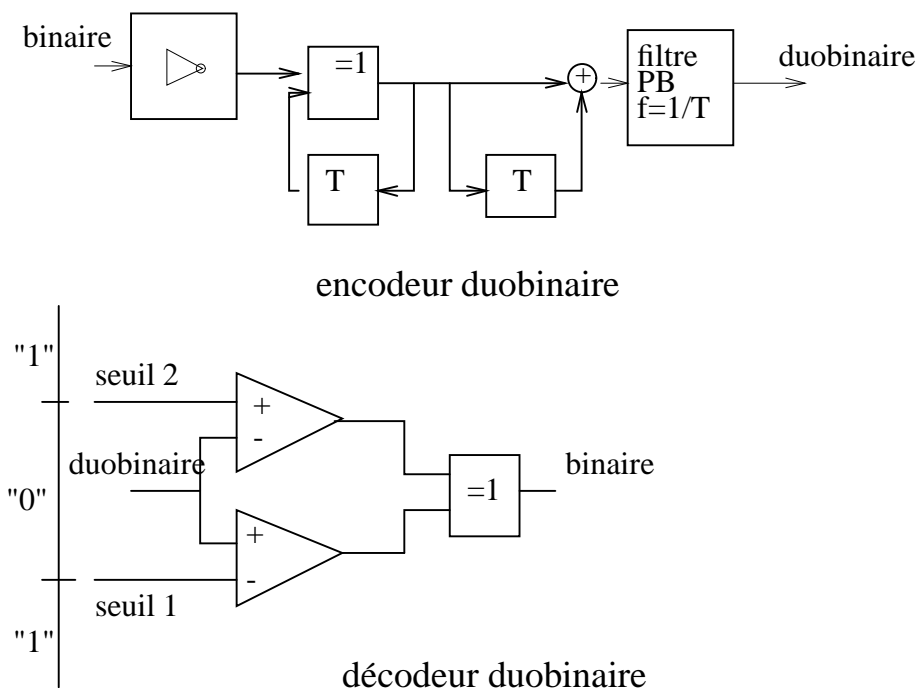


Figure 4 : codage duobinaire

D-MAC est un C-MAC dans lequel le codage binaire du son a été remplacé par un codage duobinaire (figure 4).

Le signal peut posséder un des trois niveaux électriques  $-V$ ,  $0$ ,  $+V$  correspondant aux valeurs logiques 1, 0, 1. La transmission par câble dans des canaux 10,5Mhz est possible mais elle dépasse les capacités des réseaux cablés existants (7 à 8 Mhz selon les pays). Il a donc été nécessaire de réduire la bande passante requise pour la voie son et données en n'utilisant que 4 voies mono ou 2 voies stéréo, d'où la dénomination D2-MAC.

**Trame D2-MAC** Le son et les autres informations digitales sont additionnés au signal MAC sous forme d'un paquet de 751 bits dont 720 représentent actuellement de l'information (figure 5). Les 31 bits constituent un entête et indique le

type d'information transmise dans le paquet.

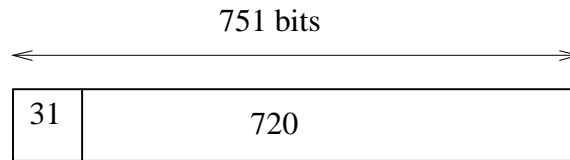


Figure 5 : paquets MAC

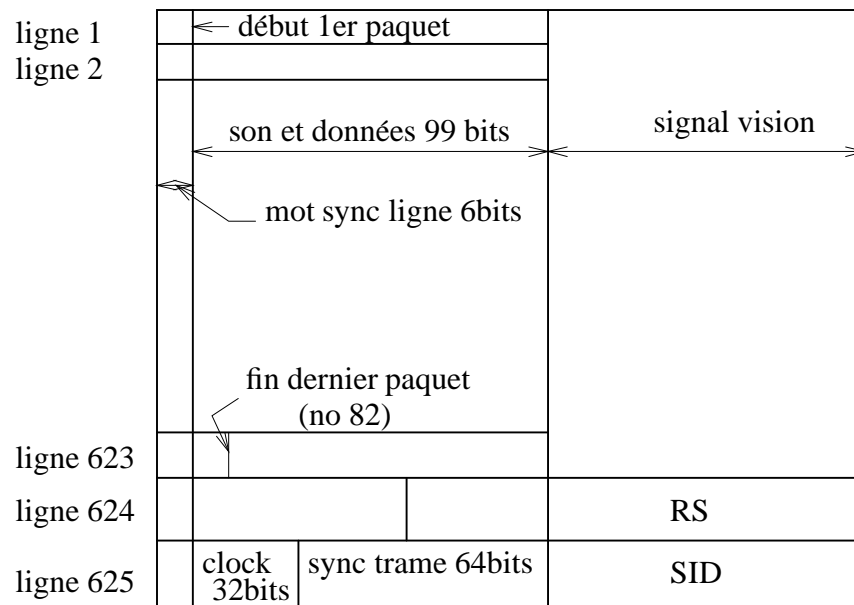
Par ligne, le groupe de données contient 105 bits utiles dont 6 bits servent à la synchronisation ligne et les 99 autres codent le son et les données.

Dans une trame (figure 6) le multiplexage son/données occupe 623 groupes et est transmis par paquets de 751 bits. Les intervalles de suppression de lignes (lignes 1 à 22) contiennent des données pour le télétexte. Il n'y a pas d'informations analogiques pour la synchronisation. Celle-ci est effectuée en évaluant le mot de synchronisation digital ligne ou le mot de synchronisation trame en ligne 625.

Le canal D.A.T.V. (Digital Assistance Television) correspond au retour de trame d'une durée équivalente à 23 lignes, soit une capacité de 1,05 Mbits/s. Il véhicule des informations exploitées par le décodeur telles que le taux de compression, le format de l'image (4:3 ou 16:9), les éventuels vecteurs de mouvement, les coefficients des filtres utilisés.

## 1.5 Avenir du D2-MAC

Le D2-MAC en tant que norme unique européenne de transmission par satellite semble compromis. En effet les différents pays n'ont pas pu se mettre d'accord sur un standard unique. On s'est ainsi retrouvé avec différentes versions de MAC. Les déboires qu'ont connus les satellites de diffusion, la non disponibilité à temps des décodeurs D2MAC, des choix politiques discutables ont sérieusement freiné la diffusion de la norme.



SID : service identification data RS : reference signals

Figure 6 : structure d'une trame D2-MAC

## 2 TVHD

Par rapport à la télévision actuelle la télévision haute définition (TVHD) se caractérise essentiellement par [5] :

- une définition verticale et horizontale améliorée dans un rapport 2 ;
- la suppression des interférences entre chrominance et luminance ;
- un format image qui passe de 4:3 à 16:9 proche de celui utilisé par le cinéma ;
- l'écran est plus grand avec une diagonale d'environ 1,5m ; des solutions transitoires avec des tailles plus classiques de 0,9m existeront ;
- un son stéréophonique avec qualité HiFi ;



## 2.1 MUSE

Les Japonais, maîtres du marché américain des équipements grand public, se sont intéressés très tôt à la TVHD. L'Europe ne s'est réveillée qu'en 1980 lorsqu'elle a réalisé que les Japonais allaient présenter au CCIR une proposition de norme mondiale pour une TVHD, 1125 lignes, 60Hz, image 5:3, entrelacée, luminance 20Mhz, chrominance bande large 7Mhz ou bande étroite 5,5Mhz, image 5:3 comme pour le cinémascope, tout en restant dans la bande des 8 Mhz. Les Japonais proposaient d'une part une norme de production d'images sur la base du système NTSC et une norme de transmission dite MUSE (*Multiple Sub-nyquist Sampling Encoding*). Ce système permet de transmettre la TVHD par compression du signal et de rester dans la limite de la bande passante de 8Mhz. Cette approche où toute la chaîne audiovisuelle depuis la production jusqu'à la visualisation était bouleversée n'assure aucune compatibilité avec le matériel existant et exige le renouvellement du parc des appareils de production et de réception.

## 2.2 HDMAC

L'Europe a réagi en proposant la TVHD selon un processus évolutif permettant le passage progressif vers la TVHD en conservant l'existant.

Le système HDMAC [2] est basé sur les acquis suivant : le système de production en composantes numériques 4-2-2 (recommandation 601 adoptée par le CCIR en 1982) et la norme de diffusion MAC . On peut ainsi transmettre, sur un canal MAC des images standards ou haute définition à destination d'un parc de récepteurs hétérogènes. Les décodeurs exploiteront, ou non selon le type de récepteur auquel on aura à faire, toutes ou partie des informations du canal DATV. Cette approche permet une évolution progressive, sans remise en cause de l'existant, vers la TVHD.

Comme il s'agit de transmettre dans un canal MAC une image haute définition il faudra utiliser des techniques de compression au niveau du codeur HDMAC et de décompression au niveau du récepteur. Le taux de compression à obtenir est de 4 pour un format d'image 4:3.

Il faut cependant conserver à l'esprit que la diffusion hertzienne pourrait à terme devenir le plus mauvais moyen pour délivrer une image de télévision couleur dans les foyers. Elle pourrait être remplacée par les fibres optiques ou d'autres systèmes

à large bande. Cela remettrait en cause les efforts de normalisation en TVHD basés sur le système MAC.

## **3 TV numérique**

### **3.1 Intérêt du numérique**

En production d'images :

- La numérisation permet de traiter l'image. On peut corriger ou modifier les couleurs, les contrastes, les contours, supprimer un élément de l'image, lui incorporer un sujet, incruster une image.
- Le numérique facilite le transcodage NTSC, PAL, SECAM, D2-MAC.

En distribution d'images :

- immunité au bruit ;
- séparation des canaux multiplexés sans diaphonie ;
- indépendance des canaux par rapport au contenu (vidéo, son, données) ;
- secret de la communication par embrouillage complexe ;

En réception :

- absence de réglage et stabilité de certains éléments du récepteur ;
- traitement de certains défauts tels que suppression d'écho, réduction de bruit, désentrelacement pour éviter le scintillement interligne, multiplication du nombre de trames pour éviter le papillotement ;

### **3.2 Recommandation CCIR 601**

Pour la production la norme numérique unique, dite norme 4-2-2 a fait en 1982 l'objet d'un accord mondial au sein du CCIR. Elle est associée à la recommandation CCIR 604. Tout en tenant compte des particularités techniques des deux grands systèmes de télévision dans le monde, ce qui a conduit à conserver deux modes d'exploitation de cette norme (625 lignes, 50 demi-images par seconde ou

525 lignes, 60 demi-images par seconde) le choix de paramètres communs fondamentaux a été fait de telle façon que les équipements numériques futurs puissent à terme fonctionner simplement et indifféremment dans l'un et l'autre mode.

Chaque composante du signal vidéo est échantillonnée sur 8 bits à la fréquence de 13,5 Mhz pour la luminance et 6,75 Mhz pour la chrominance. Ceux-ci sont transmis selon la séquence 4 échantillons de luminance suivis de 2\*2 échantillons de chrominance (2 Cr et 2 Cb).

### 3.3 Débits

<i>type image</i>	<i>HDP</i>	<i>HDI</i>	<i>EDP</i>	<i>4:2:2 (4/3)</i>	<i>4:2:2 (16/9)</i>	<i>VT (4/3)</i>	<i>VT (16/9)</i>
f frame	50	50	50	50	50	50	50
lignes	1152	1152	576	576	576	288	288
pixels	1920	1920	960	720	960	360	480
balayage							
P progressif	P	I	P	I	I	P	P
I entrelacé							
débits Mbits/s	1700	885	442	166	221	83	111

HDP : high definition progressive, HDI : high definition interlaced

EDP : enhanced definition progressive, VT : videotelephone.

## 4 Supports de transmission de l'image

### 4.1 Supports physiques

L'image peut être transmise par voie hertzienne, sur un réseau ou sur une liaison locale, à faible ou courte distance [13].

- Câble coaxial.

Comme les pertes de transmission deviennent importantes quand la fréquence augmente, il est nécessaire d'utiliser des répéteurs. Il en résulte que le câble coaxial ne peut être utilisé sans technique de compression et son usage est réservé pour des transmissions de l'ordre de 100-200 mètres et non pour un usage de type télécommunication.

- Fibre optique.  
Les caractéristiques de la fibre optique, large bande passante et faible perte de transmission, en font un support de transmission économique.
- Diffusion par satellite.  
La WARC en 1977 a attribué à la diffusion par satellite la bande de fréquence comprise entre 11,7 et 12,5 Ghz. Chaque pays européen dispose de cinq bandes (BP=27Mhz). La largeur de bande d'un canal est de 27Mhz et l'espacement entre deux canaux est de 19,18Mhz. Deux canaux adjacents sont émis avec des polarisations croisées (protection). Les caractéristiques d'un signal D2-HDMAC/Paquet sont compatibles avec la transmission en modulation de fréquence dans un canal de 27Mhz.  
Normalement pour les communications par satellite un transpondeur avec une bande passante de l'ordre de 100 Mhz serait nécessaire pour transmettre de la TVHD en modulation de fréquence. Pour cet usage il est pratique et économique d'utiliser des transpondeurs à usage général qui sont normalement utilisés pour transmettre des signaux TV conventionnels dont la bande passante est comprise entre 27 Mhz et 36 Mhz. Il est donc nécessaire de réduire la bande passante du signal ou d'utiliser un codeur-décodeur efficace.
- Communications terrestres.  
La bande passante est aussi de 40Mhz au maximum.
- Transmissions à courte distance.  
La transmission par fibre optique est la plus économique. Pour les transmissions analogiques on peut envisager de transmettre les trois composantes du signal vidéo couleur en multiplex de fréquence ou temporel.
- Communications terrestres à longue distance.  
Comme les coûts de transmission dépendent du câble et des répéteurs utilisés, un système de transmission numérique multiplexé tel que celui utilisé pour le téléphone sera utilisé pour la transmission du signal TVHD. Le débit doit correspondre à la hiérarchie de fréquence utilisée par le téléphone. Comme le coût est proportionnel au nombre de canaux téléphoniques utilisés il est économique de mettre en œuvre des techniques de compression.

## 4.2 Contribution et distribution

En matière de transmission audiovisuelle le CCIR définit des applications de contribution et de distribution par des critères de qualité distincts. Une liaison de contribution est prévue, dans le cadre d'un usage professionnel, pour permettre un traitement ultérieur du signal transmis. Il concerne des liaisons de studio à studio. Une liaison de distribution est orientée utilisateur et ne nécessite pas forcément la qualité précédente. Il concerne les liaisons de studio à station de diffusion (distribution primaire) ou entre celle-ci et l'utilisateur (distribution secondaire). Il est possible d'avoir une distribution primaire digitale et une distribution secondaire analogique.

## 4.3 Hiérarchie de débits

Actuellement les débits des canaux disponibles, en France, pour les réseaux publics (ISDN) sont de 2,048 Mbits/s, 34,368 Mbits/s et 139,264 Mbits/s conformément à la recommandation G702 du CCITT (figure 7).

La S.D.H. (Synchronous Digital Hierarchy) a été recommandée par le CCITT en 1988. Elle est basée sur des multiples de  $D = 155,52$  Mbits/s et propose des canaux à  $n \cdot D$  Mbits/s jusqu'à 2,488 Gbits/s ( $n=16$ ) [12].

Sous SDH des connexions internationales peuvent être établies économiquement.

## 4.4 Supports et applications

Les algorithmes de codage, choisis parmi les recommandations existantes ou à venir, vont dépendre des applications visées ainsi que des supports physiques et des canaux utilisés.

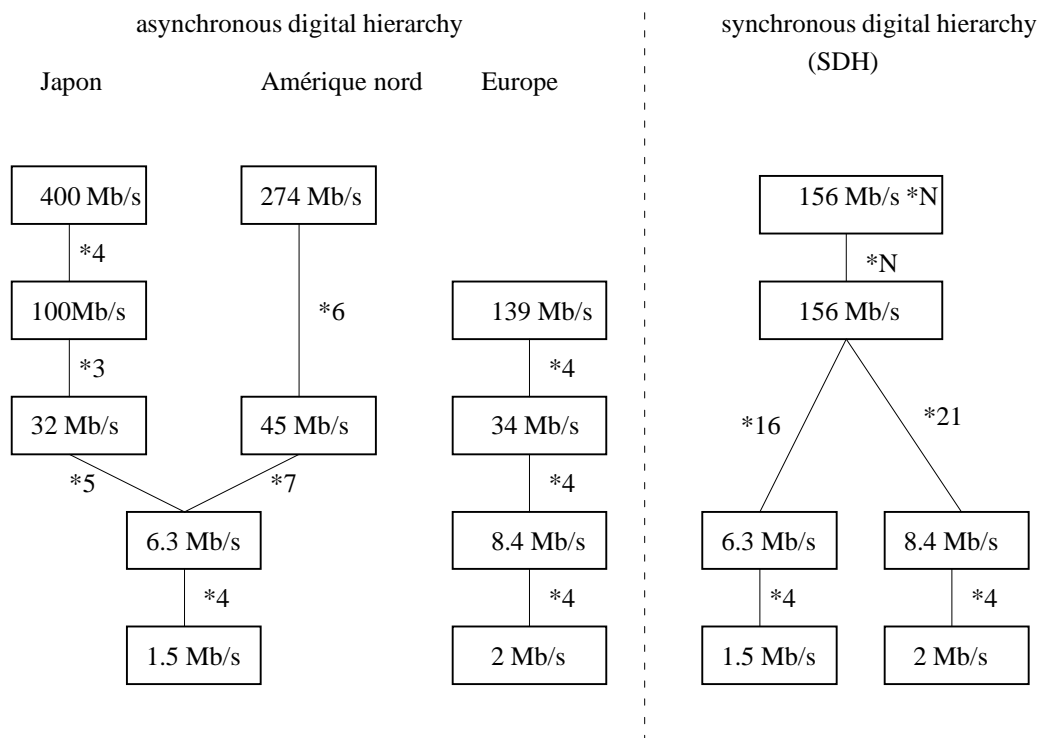


Figure 7 : caractéristiques des canaux sur réseau

#### 4.4.1 Réseaux

[10]

<i>canal ISDN</i>	<i>débits</i>	<i>applications</i>	<i>codage</i>	
B	64 Kbps	audiovisuel	H261	
2B	128 Kbps	audiovisuel	H261	
..	..	..	..	
6B	384 Kbps	audiovisuel	H261	
H0	384 Kbps	audiovisuel	H261	
..	..	..	..	
H11	1,536 Mbps	audiovisuel	H261	
H12	1,92 Mbps	audiovisuel	H261	
H21	30,720 Mbps	diffusion	33Mbits	
H22	33,792 Mbps			
H23	44,160 Mbps	contribution		
H3	≈ 70 Mbps			
H4	≈ 140 Mbps			

#### 4.4.2 Applications audiovisuelles

Les applications audiovisuelles telles que le vidéophone et la visioconférence utilisent les canaux B à H12 et la standardisation est à la charge du groupe d'étude XV du CCITT. Celui-ci a émis des propositions de recommandation pour adoption par le CCITT afin d'harmoniser les systèmes apparaissant sur le marché suite au développement des circuits à 64 Kbits pour les réseaux RNIS. Cette harmonisation vise aussi bien le matériel, que les protocoles de communication et les algorithmes de codage. Il a émis la recommandation H261 pour le codage vidéo à p\*64Kbits pour les applications vidéotéléphone et vidéoconférence de 64 Kbits/s à 2 Mbits/s.

#### 4.4.3 Diffusion TV

L'étude de standardisation de réduction de débit par codage pour les débits de 30 à 140 Mbits/s et plus particulièrement pour la plage 30-34 Mbits/s [8] entreprise par le CMTT a abouti à la recommandation 723 [4].

Le canal H4 (140 Mbits/s) est le débit le plus prometteur pour transmettre des signaux entre centres de production pour des applications de contribution. Le projet Race-1018 a réalisé un codeur permettant de transmettre des images HDI sur un canal H4 [12]. Le taux de compression à atteindre est de 2,2 bits/pel.

Le canal H2 (33 Mbits/s) est le plus adapté à la transmission longue distance, à faible coût. Il s'agira donc de passer d'un débit de 166 Mbit/s à 33 Mbits/s (contribution) ou 15 Mbits/s (distribution). Ceci correspond à une compression de 16 bits/pel (luminance et chrominance) à 2,9 bits/pel et 1,45 bits/pel respectivement. A cet effet le CCIR a retenu l'algorithme H261.

#### 4.5 Voie hertzienne

Il s'agit de diffuser par liaison satellite ou liaison terrestre un signal HDMAC sur un canal MAC (bande passante 8,4 Mhz)

### 5 Etudes de compression

De nombreux projets et groupes de travail, dans le cadre d'instances internationales, ont étudié les différents maillons de la chaîne image, depuis la production (caméra) jusqu'à la visualisation (récepteur) en passant par les codeur-décodeurs nécessaires pour adapter les débits images au support de transmission choisi. Les

domaines d'applications concernés vont de l'image fixe à la télévision haute définition. Ces études ont débouché pour certaines sur des recommandations, pour d'autres l'évaluation de diverses approches est encore en cours.

Trois recommandations concernant la compression, dépendant du type d'application, ont été développées pour des images fixes (JPEG), mobiles (MPEG) et la vidéoconférence (H261). Ceci permet aux industriels produisant les différents maillons de la chaîne de proposer des solutions compatibles.

Un effort identique de normalisation est en cours pour le domaine de la télévision où il s'agit de préparer l'avènement de la haute définition dans un esprit de compatibilité avec le parc existant (HDMAC) et d'harmoniser les échanges d'images pour des applications de contribution ou de diffusion (33 Mbits/s).

## **5.1 Quelques projets**

### **5.1.1 RACE**

Les projets de la communauté européenne faisant partie de RACE sont dédiés aux communications entièrement numériques incluant les services de diffusion TV et HDTV (RNIS à très large bande). Les solutions proposées dans le cadre de RACE s'appuient largement sur des systèmes commutés totalement numérisés et utilisant de la fibre optique jusqu'en distribution finale ; c'est une extrapolation des réseaux téléphoniques. On peut légitimement s'interroger sur la validité des choix de RACE en particulier sur la nécessaire compatibilité des réseaux de télédistribution avec la TVHD et avec les services bidirectionnels. Tout se passe comme si RACE avait pris conscience des limites commerciales des réseaux multiservices en étoile commuté et se repliait sur le marché de la télévision distribuée par câble qui peut utiliser les mêmes lignes de distribution en fibre optique mais qui peut se contenter d'une commutation très simplifiée.

Le projet Race-1018 vise à réaliser un codeur/décodeur pour des liaisons de contribution, utilisant l'algorithme du CCIR. Il permettra de transmettre des images au format HDP pour la contribution sur un canal 140 Mbits/G702 ou 155 Mbits/SDH. Le taux de compression à atteindre est de 2,2 bits/pel.

### **5.1.2 EUREKA95**

EUREKA95 est une réponse au projet japonais, il a pour but de faire la démonstration qu'une norme de TVHD évolutive, compatible et s'appuyant sur des normes



de transmission de la famille paquet était possible. EUREKA95 est divisé en 10 sous-groupes couvrant l'ensemble de la chaîne de l'image : caméra, magnétoscope, transmission, visualisation.

La norme de production unique retenue par EUREKA95 est la norme à 1250 lignes, 50Hz progressif ou HDP à 2,4Gb/s. Cependant en attendant cette norme unique, des solutions intermédiaires ont été proposées.

- HDQ : système issu du HDP avec un filtrage diagonale permettant d'en réduire le débit à 1,2 Gbits/s.
- HDI : système entrelacé à 1250 lignes.

La transmission s'appuie largement sur le HDMAC. Il n'y a pas toujours cohérence technique entre les projets EUREKA et RACE, les échéances n'étant pas les mêmes et les industriels et laboratoires souvent différents.

Le codeur-décodeur développé dans le cadre du projet Eureka 256 met en œuvre une technique de compression à base de DCT, prédiction par compensation de mouvement et codage VLC pour des fréquences de transmission de 45 à 140 Mbits/s. La RAI dispose d'une dizaine de codeurs prototypes.

JESSI est un programme EUREKA destiné à mettre à niveau les industriels européens (Siemens, Philips, SGS Thomson) en matière de mémoire à haute performance et relever ainsi le défi japonais. Ce programme devrait permettre d'obtenir les mémoires 16Mbits des récepteurs TVHD.

### 5.1.3 TVHD américaine

L'ATTC (Advanced Television Testing Center) est chargée d'examiner les prétendants au futur standard américain [1] et de fournir à la FCC (Federal Communications Commission) les résultats de ces tests en vue de l'adoption d'un système de TVHD pour la fin du second semestre 1993. De toute façon la solution retenue sera de type *simulcast* compatible NTSC.

Parmi les propositions on trouve :

- Le système DigiCipher de General Electric (définition de 1050 lignes), développé en collaboration avec le MIT, entièrement numérique, s'appuyant au niveau de la compression de données sur des algorithmes propriétaires (compensation de mouvement estimée par superblocs, association de plusieurs blocs de DCT).

- Le système Advanced Compatible Television de l'ATCR, partiellement analogique (définition de 1050 lignes). Compression Labs apporte son soutien à l'EDTV (Extended Definition Television), proposée par l'ATRC et mise au point par le laboratoire David Sarnoff. L'ATRC est composé du centre de recherche David Sarnoff, de Philips Consumer Electronics, de Thomson Consumer Electronics et de NBC. L'EDTV est une étape intermédiaire entre la télévision actuelle et le système de TVHD tout numérique ADTV (Advanced Digital Television). Dans ce cas un challenge serait d'utiliser les canaux de 6 Mhz, affectés à la transmission hertzienne des programmes NTSC.
- Le système Digital Spectrum Compatible de Zenith et AT&T, tout numérique (définition de 1575 lignes).
- Le Narrow Muse du japonais NHK (définition de 1125 lignes).

## 5.2 Algorithmes normalisés

Ces projets et d'autres ont abouti à la définition de normes ou de recommandations par les instances internationales et ceci pour un large spectre d'applications. Tous les aspects de la chaîne image sont bien entendus concernés, mais l'on ne s'intéressera ici qu'à la partie algorithme de compression.

### 5.2.1 H 261

L'algorithme CCIR H261 a été retenu pour les applications vidéotéléphone/vidéconférence de 64 Kbits/s à 2 Mbits/s et par le CCIR pour les applications TV de contribution à 34 Mbits/s. Dans le cas du visiophone pour un codage à 64kbits/s au format QCIF 10 Hz, le nombre de bits disponible par point est de 0,16.

La recommandation H221 indique comment sont multiplexées et synchronisées les informations (parole, données, images) sur plusieurs canaux 64kbits/s utilisés par un terminal. Par exemple un terminal peut utiliser 16 kbits/s de données audio et 110 kbits/s de données vidéo sur un canal 2B (128 kbits/s). Les images NTSC/PAL/SECAM sont transformées en format CIF (*Common Intermediate Format*) qui est le format d'entrée du codeur H261. Le format CIF permet de prendre en compte les problèmes d'interfonctionnement entre régions

du monde qui possèdent des standards de télévision différents

Les caractéristiques de CIF sont :

- nombre de lignes par image :  $Y=288$ ,  $Cr=144$ ,  $Cb=144$  ;
- pixels par ligne :  $Y=360$ ,  $Cr=180$ ,  $Cb=180$  ;
- fréquence image 29,97 Hz ;
- mode non entrelacé ;

Un format QCIF (*Quater CIF*) a aussi été défini qui correspond à une image de résolution moitié dans les deux directions.

### 5.2.2 VISAGE

La transmission des images animées, pour des applications de type visiophone, requiert des capacités sans commune mesure avec celles du téléphone (débits 1000 fois plus importants) et la visiophonie nécessiterait l'installation d'un réseau commuté large bande. On ne peut guère retenir cette solution à court terme et l'arrivée du RNIS (réseau commuté offrant deux voies à 64kbits/s pour l'accès de base) a conduit à étudier une autre approche. Les progrès en matière de compression permettent d'envisager de transmettre des images animées de type visiophonique (personnage en buste) sur une voie de 64kbits/s. Le projet VISAGE, démarré en 1987, est un projet de visiophone sur RNIS : un canal est utilisé pour le son l'autre pour l'image. Ce projet a contribué à l'élaboration de l'algorithme à p\*64kbits/s et la deuxième génération de codeurs sera conforme à la recommandation H.261 du CCITT avec des formats d'image CIF.

### 5.2.3 HDMAC

Le support de transmission retenu, dans l'esprit de la compatibilité, est un canal MAC qui permet la diffusion d'images de définition normale (625 lignes, 720 points). La transmission d'images haute définition de 1250 lignes de 1440 points sur un canal MAC nécessite une compression de la bande du signal par un facteur 4.

Le signal HDMAC ainsi compressé peut, après transmission sur un canal MAC, être visualisé sur un récepteur standard possédant un décodeur MAC ou visualisé sur un récepteur haute définition après décodage HDMAC.

**Codeur à 3 branches** Le principe de compression de bande retenue (HDMAC) [9] consiste à adapter la structure d'échantillonnage spatio-temporelle au mouvement détecté sur chaque bloc 8\*8 de l'image. On obtient ainsi un codeur à 3 branches (figure 8).

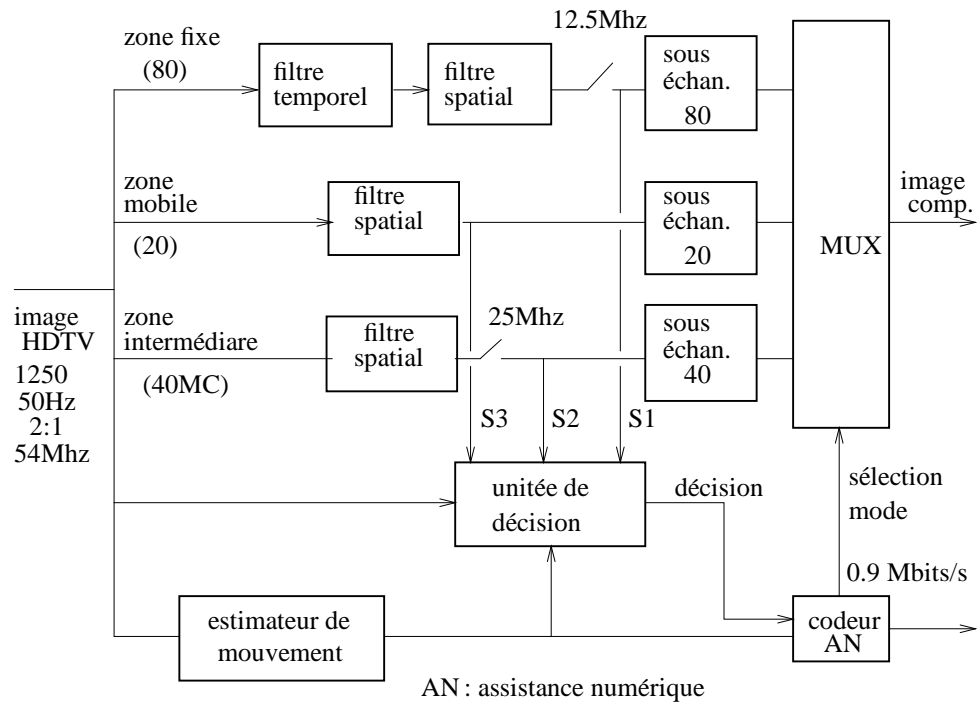


Figure 8 : codeur HDMAC

On rappelle qu'une image est transmise sur deux trames (mode entrelacé) à raison d'une trame toutes les 20ms.

La compression d'un facteur 4 est obtenue ainsi :

- Blocs stationnaires ( $0 \leq \text{mouvement} \leq 0,5 \text{ pixels}/40\text{ms}$ ); la transmission de l'image (N pixels) se fait sur 80ms (mode 80). On transmet une image sur deux; les images retenues sont sous-échantillonnées dans un rapport 2 et transmises sur 4 périodes de 20ms (N/8 pixels). On obtient ainsi la plus forte résolution spatiale.

- Blocs intermédiaires ( $0,5 \leq \text{mouvement} \leq 1,5 \text{ pixels/40ms}$ ); on transmet une trame sur 2 ; les trames retenues sont échantillonnées dans un rapport 2 et transmises chacune sur deux périodes de 20ms (mode 40). La résolution spatiale et temporelle sont divisées par 2.
- Blocs mobiles ( $\text{mouvement} \geq 1,5 \text{ pixels/40ms}$ ); chaque trame est transmise avec un sous-échantillonnage de 4. On conserve la totalité de la résolution temporelle au détriment d'une résolution spatiale dégradée dans un rapport 4 (mode 20).

Les structures de sous-échantillonnage utilisées sont les suivants :

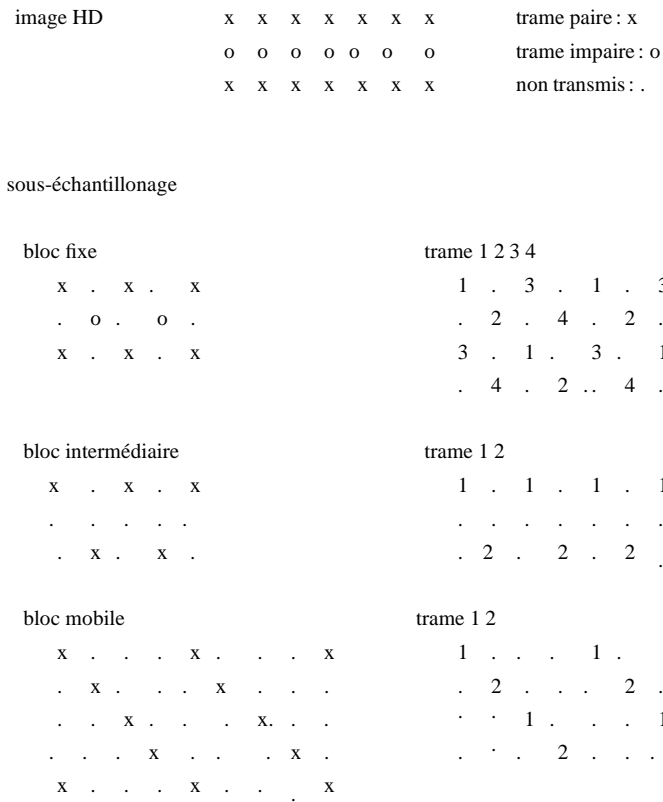


Figure 9 : sous échantillonnage HDMAC

Pour les bloc stationnaires et mobiles un brassage de lignes (figure 10) est nécessaire pour constituer les trames finales (durée ligne =  $64\mu s$ ) à partir des trames sous-échantillonnées.

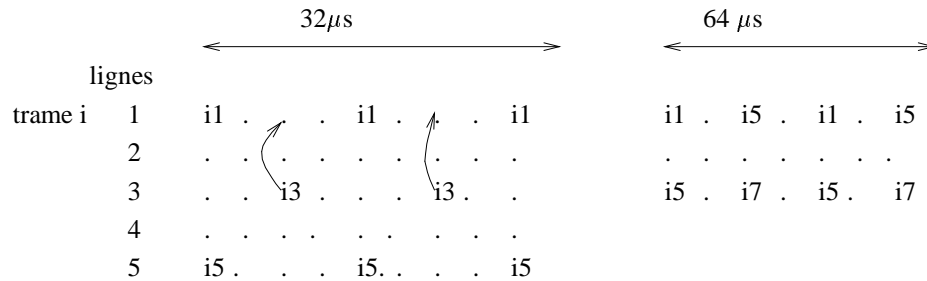


Figure 10 : brassage de lignes

Ainsi dans le cas de mouvements de faible amplitude, la transmission d'une image se fait sur quatre trames. On obtient alors une définition spatiale 4 fois supérieure à celle obtenue dans le cas de mouvements importants où la transmission d'une image s'effectue en deux trames entrelacées. Dans ce dernier cas toute la résolution temporelle est conservée au détriment de la résolution spatiale. Entre ces deux modes, le mode intermédiaire emploie le principe de la compensation de mouvement afin d'obtenir une reconstitution satisfaisante. L'image étant sous-échantillonnée au niveau du codeur puis sur-échantillonnée au niveau du décodeur, il est nécessaire de prévoir des filtres aux deux extrémités de la chaîne de traitement pour assurer une reconstitution de l'image sans recouvrement spectral.

**Sous-branches** le sous-échantillonnage décrit précédemment est du type en quinconce. Il y a des types d'images pour lesquels ces filtres ne sont pas les mieux adaptés. Pour cela il a été proposé [11] d'introduire des sous-branches permettant de choisir un type de filtre pour chaque type de blocs. Le choix des sous-branches peut se faire *a priori* ou *a posteriori*.

**Estimation de mouvement** Il s'agit d'estimer le mouvement de chaque bloc  $16*16$  pour d'une part déterminer quelle branche du codeur devra être sélectionnée

et d'autre part calculer un vecteur de mouvement pour la trame non transmise, en mode 40, afin que le décodeur puisse la reconstituer à partir de la trame précédente.

**Signal d'assistance numérique** Dans ce système de transmission à 3 modes il est nécessaire de transmettre une information relative au mode choisi ainsi que des informations de mouvement qui permettent la reconstruction des images haute définition à réception. Dans le standard MAC la période de retour trame représente un canal disponible d'une capacité d'environ 1 Mbits/s et permet, moyennant un système de réduction de débit, la transmission de ces informations dites d'assistance numérique (vecteurs de mouvement, coefficients de filtres, format image, type de séquence de modes)

Dans le système HD-MAC retenu par Eureka, la commutation entre les 3 modes s'effectue sur la base de blocs de 16 lignes de 16 échantillons pris dans l'image HD de 1250 lignes de 1440 pixels. Le mode choisi est celui qui permet de minimiser l'énergie de la différence entre l'image originale et l'image reconstruite. Pour chacun des blocs il faut transmettre le type de filtrage utilisé. Dans le cas du mode intermédiaire un vecteur de mouvement associé au bloc est nécessaire. La manière dont l'information mouvement est transmise fait l'objet de normalisation. Par contre la manière de calculer ce vecteur mouvement est arbitraire.

**Commutation entre mode** On assure une cohérence spatiale entre les blocs. Un bloc en mode 80 isolé dans un voisinage de blocs en mode 20 verra son mode forcé à 20.

Pour éviter les ruptures trop fréquentes de phases d'échantillonnage, une cohérence temporelle sur 4 trames est introduite dans la prise de décision. Cette cohérence limite à cinq les différents séquences possibles de modes sur un intervalle de 80ms : (mode 80), (40,40), (40, 20, 20), (20, 20, 40), (20, 20, 20, 20). Par ailleurs une cohérence temporelle est aussi assurée pour une succession de séquences [6].

#### 5.2.4 JPEG

Le standard JPEG (Joint Photographic Expert Group [14]) concerne des images de dimensions, caractéristiques de couleurs, de précision de pixels quelconques, présentant une continuité temporelle (imagerie médicale, art, photo de presse ). Elle s'adresse principalement aux télécopieurs, vidéotexte, imprimantes et appareils photos numériques. JPEG a pour objectif de compresser des fichiers d'images

avec un taux de compression sélectionnable par l'utilisateur.

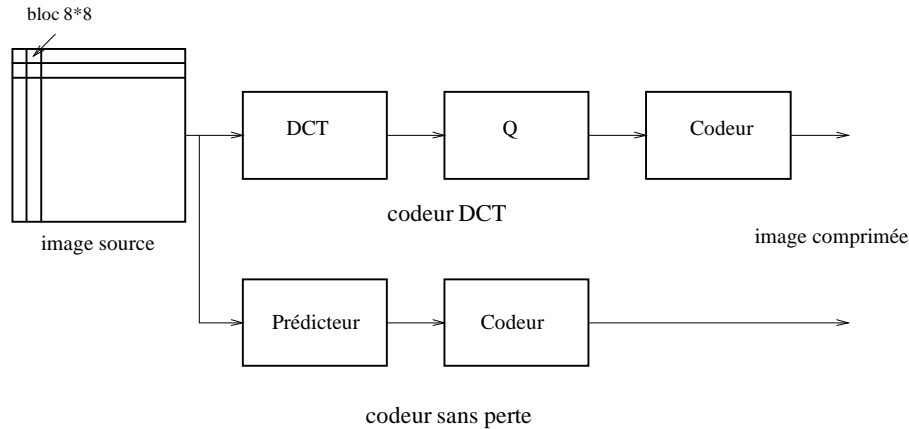


Figure 11 : variantes de codeurs JPEG

**Modes** Les modes opératoires suivant sont supportés :

- Codage séquentiel : chaque composante de l'image est codée par balayage gauche-droite et haut-bas.
- Codage progressif : chaque composante de l'image est codée en plusieurs passes. La première passe est grossière et permet une transmission rapide par rapport au temps d'émission disponible. Cette première version peut être affinée par des passes successives jusqu'à atteindre la qualité d'image souhaitée déterminée par la table de quantification. On peut ainsi au cours des passes successives générer les coefficients successifs ou affiner la précision de tous les coefficients.
- Codage hiérarchique : l'image est codée à plusieurs niveaux de résolution de manière à disposer de la résolution la plus faible sans avoir à décompresser l'image à sa résolution maximum. Chaque résolution diffère de la précédente d'un rapport 2 dans les 2 dimensions.



- Codage sans perte : (différentiel, sans DCT ni quantification).

Le standard proposé décrit les 4 modes précédents et spécifie un codeur pour chacun d'eux.

Le codeur JPEG (figure 11), dans le cas d'une image à une seule composante, se compose d'une DCT 8\*8 (remplacée par un prédicteur pour la version sans perte), d'un quantifieur (absent dans la version sans perte) et d'un codeur entropique (codage de Huffman ou codage arithmétique). Il n'y a pas de réduction de redondance temporelle. La composante continue est quantifiée en différentielle et les autres composantes de la DCT sont balayées en zig-zag.

Le codeur sans perte est constitué d'un prédicteur, qui peut être sélectionné parmi plusieurs combinaisons linéaires de pixels voisins et d'un codeur entropique.

**Modèle d'image** Une image source peut contenir 1 à 255 composantes (aussi appelées bandes ou canaux spectraux ou de couleur). Chaque composante consiste en un tableau d'échantillons. Un échantillon est défini par un entier codé sur  $p$  bits soit 8 ou 12 bits pour les codeur-décodeurs basés sur la DCT et 2 à 16 pour le codage prédictif. La précision est identique pour toutes les composantes. Selon l'échantillonnage des composantes, pas forcément identique, les composantes n'ont pas la même dimension. Les blocs des différentes composantes sont émis de manière entrelacée en paquets appelés MCU (*Minimum Coded Unit*). Dans ce cas des tables de quantification différentes peuvent être associées à chaque composante.

**Taux de compression** Les taux de compression donnés concernent des images couleurs moyennement complexes et l'unité bit/pixel inclue les informations de luminance. Les taux de compression spécifiés sont compris entre 10:1 et 200:1. Il dépendent de la qualité requise pour l'image décompressée. Jusqu'à 40:1 la différence est pratiquement invisible à l'oeil nu. Ces niveaux de qualité qualitatifs sont donnés à titre d'information, sachant que la qualité et la compression peuvent varier de façon significative suivant les caractéristiques de l'image et le contenu de la scène.

- 0,25 à 0,5 bits/pixels : assez bonne à bonne qualité, suffisante pour certaines applications.

- 0,5 à 0,75 bits/pixels : bonne à très bonne qualité, suffisantes pour beaucoup d'applications.
- 0,75 à 1,5 bits/pixels : excellente qualité, suffisante pour la plupart des applications.
- 1,5 à 2,0 bits/pixels : non distinguable de l'original.

### 5.2.5 34 Mbits/s

L'algorithme de codage hybride retenu par le CCITT pour la norme H261 a aussi été retenu par le CCIR pour les applications de contribution de 32 à 45 Mbits/s. L'image source comporte 288 lignes par trame entrelacée de 720 échantillons de luminance (360 pour la chrominance) codés au format 4:2:2.

Pour chaque macro-bloc un des trois modes peut être sélectionné : intra-trame, inter-trame et inter-image avec compensation de mouvement. Le taux de compression à atteindre est de l'ordre de 5.

### 5.2.6 MPEG

La proposition de standard MPEG (Moving Picture Expert Group) proposée par le groupe de travail WG11 (ISO/IEC WG11 de /JTC1/SC2) formé par l'ISO et le CCITT s'intéresse à la compression vidéo (de l'ordre de 1,5 Mbits/s) et audio (64, 128 et 198 Kbits/s) d'images animées pour des applications de stockage d'images sur des supports comme le CD-ROM, DAT (Digital Audio Tape) et les supports de stockage informatiques [7]. Les techniques mises en œuvre à cette occasion devraient pouvoir être utilisées pour les communications par réseau ISDN ou même pour des applications de diffusion. A des débits de 1,2 Mbits/s on a obtenu une bonne qualité d'images (20 à 30 images/s, 360 pixels par ligne). La compression MPEG permet d'obtenir des images de qualité VHS à partir d'un disque compact. Le taux de compression est alors de 50:1

**Particularités** Les spécifications des systèmes MPEG indiquent comment sont combinés et accédés plusieurs flux vidéo et audio regroupés en un seul flux de données. Cela permet la recherche rapide avant et arrière.

Pour la vidéo il est indispensable de pouvoir accéder de manière aléatoire (codage intra-image) à un enregistrement compressé et ceci en un minimum de temps,

de l'ordre de la demi seconde. La difficulté dans la spécification de l'algorithme MPEG est la suivante : d'une part on exige une haute qualité à un fort taux de compression que l'on ne peut obtenir par un codage intra-image; d'autre part l'exigence d'accès aléatoire est mieux satisfaite avec le codage intra-image. L'algorithme ne peut satisfaire ces exigences contradictoires qu'en réalisant le fort taux de compression par du codage inter-image sans compromettre l'accès aléatoire. Ceci nécessite un équilibre entre codage inter et intra, ainsi qu'entre réduction de redondance temporelle réursive et non réursive. Ceci a été résolu en utilisant deux techniques de codage inter-image : prédictive et interpolative.

**Codeur** Le schéma fonctionnel du codeur H261 est utilisé. Par contre les caractéristiques de l'estimation de mouvement, du quantifieur et du codeur sont différentes. On utilise donc la compensation de mouvement par bloc pour la réduction de la redondance temporelle et la DCT pour la réduction de la redondance spatiale. La technique de compensation de mouvement est appliquée à l'aide de prédicteur soit causal (codage prédictif pur) soit non causal (codage interpolatif). Le signal d'erreur est compressé avec une DCT. L'information relative au mouvement est basée sur des blocs 16\*16 et transmise avec l'information spatiale. L'information mouvement est aussi compressée en utilisant des codes VLC.

**Format image** Le format d'entrée de MPEG est appelé SIF (*Source Input Format*). C'est un format CCIRR-601 sous-échantillonné par deux horizontalement et par deux temporellement pour la luminance, à quoi il faut ajouter un sous échantillonnage vertical de deux pour la chrominance. Le format d'image est 4:3. Il existe deux variantes du format SIF : SIF 525 (60Hz) et SIF 625 (50Hz). Le SIF 625 correspond donc à 288 lignes (144 lignes) de 360 (180) échantillons pour la luminance (chrominance).

**Réduction de redondance temporelle** A cause de l'importance de l'accès aléatoire trois types d'images sont considérées dans MPEG :

- Image intra (I) : ce sont les points d'entrée pour l'accès aléatoire mais avec un faible taux de compression.

- Images prédites (P) : elles sont codées avec référence à une image précédente (de type I ou P).
- Images interpolées (B) (prédiction bidirectionnelle) : elles nécessitent une image passée et une image future pour la prédiction, mais ne sont jamais utilisées comme image de référence.

Une structure répétitive d'image I,B,P constitue un groupe d'image (figure 12). Le rapport entre le nombre d'images I et P est de 3 à 4.

Pour les macro-blocs (16\*16) des images B il y a quatre types de prédicteurs possibles : intra, avant, arrière et moyen.

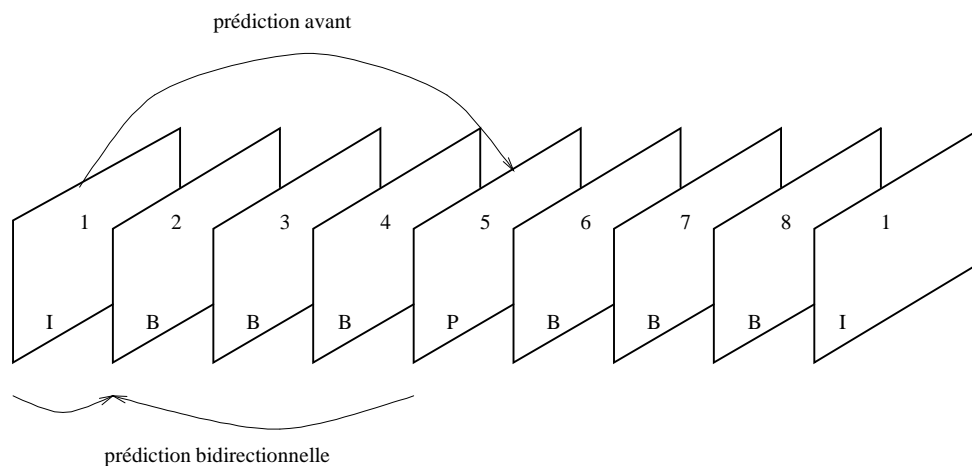


Figure 12 : groupe d'images MPEG

**Quantification** Les blocs codés intra contiennent toutes les fréquences et l'effet de bloc apparaît facilement si la quantification est trop grossière. Pour les blocs codés en différentiel on trouve essentiellement des composantes hautes fréquences en supposant que le codeur sache prédire précisément les faibles fréquences. Le pas de quantification peut être adapté aux caractéristiques d'un bloc.

**Futur** Un nouveau standard de compression pour des images de plus grande résolution, correspondant à des débits de 5 à 10 Mbits, est en cours d'élaboration (MPEG2). Les applications visées sont le stockage d'images et la diffusion sur réseau.

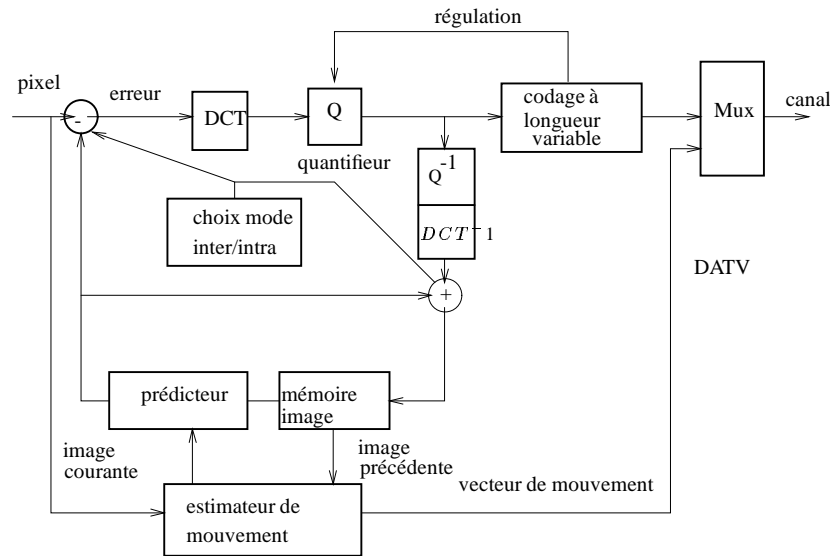


Figure 13 : codeur H261

## 6 Quelques sigles

- **ATV : Advanced Television.** A system that distributes wide-screen television signals with resolution substantially improved over 525-line and 625-line systems. Terrestrial ATV broadcasting (VHF/UHF) would require substantial changes to current emission regulations. There is general acknowledgement that any ATV distribution system should offer at least stereo (two channel) audio, of CD quality
- **ATV : Transmission.** The distribution of ATV program material to the ultimate viewing audience through RF media such as terrestrial VHF/UHF

broadcasting, main channel. It is a technical challenge in an augmentation system to avoid visible seams between side panels and the main picture

- **Simulcast ATV** : A system that transmits a self-contained ATV signal and achieves receiver-compatibility by simultaneously transmitting the same program material, possibly with reduced aspect ratio, in a separate 525/59.94 or 625/50 channel. Zenith SC-HDTV is an example of such a system that in this case is optimized to exploit a currently-unused VHF/UHF taboo channel to convey the self-contained ATV signal. Sometimes confused with Incompatible ATV cable television (CATV) or direct broadcast satellite (DBS).
- **Augmentation ATV** : An ATV transmission system that transmits an augmentation signal that is intimately associated with a main channel signal conformant to current broadcast standards. The two signals can be combined in an ATV receiver to recreate a complete ATV signal. Augmentation systems provide receiver-compatibility by default. No 625-line country is currently planning to use an augmentation channel approach for ATV
- **In-Band Augmentation ATV** : An ATV transmission system that alters the 525-line or 625-line broadcasting standard by the addition of an augmentation signal within the RF channel bandwidth of the current standard. Such systems necessarily degrade reception of current signals. Sarnoff ACTV is an example of an in-band augmentation ATV system
- **Out-of-Band Augmentation ATV** : A system that transmits an augmentation signal in an RF channel outside the main channel. The NYIT VISTA system is an example of an out-of-band augmentation system
- **Compatible ATV** : An ATV transmission system that is intimately related in some manner to current broadcast. Although it is a self-contained ATV system, Zenith SC-HDTV is called compatible because of its use of a frame rate of 59.94 Hz, its line rate of three times current 525/59.94 and the relationship of its RF channel utilization with NTSC
- **BISDN** : Broadband Integrated Services Digital Network, réseau de communication basé sur les fibres optiques

- CCIR : Comité Consultatif International pour les Radiocommunications, organe des Nations Unies, qui par voie de recommandations entérine les normes internationales
- CDR : Common Data Rate. The standardization of a single data rate for digital HDTV, independent of the choice of image format, frame rate and sync/blanking structure
- CEPT : Commission Européenne des Postes et Télécommunication
- CFR : Common Frame Rate. The standardization of a single frame rate for digital HDTV.
- CIF : Common Image Format. The standardization of the structure of the samples that represent the picture information of a single frame in digital HDTV, independent of frame rate and sync/blanking structure.
- CMTT : Committee for Mixed Telephone and Television (joint CCIR/CCITT)
- EACEM : European Association of Consumer Electronics Manufacturers association de constructeurs de matériel grand public européen
- EDTV : Extended [or Enhanced] Definition Television. A television system that offers picture quality substantially improved over conventional 525-line or 625-line receivers, by employing techniques at the transmitter and at the receiver that are transparent to (and cause no visible quality degradation to) existing 525-line or 625-line receivers. One example of EDTV is the improved separation of luminance and colour components by pre-combing the signals prior to transmission, using techniques that have been suggested by Faroudja, Central Dynamics and Dr William Glenn
- ESAC : Economics and Statistics Advisory Committee
- ESPRIT : European Strategic Programme for Research and Development in Information Technology )
- ETSI : European Telecommunication Standard Institute, organisme de normalisation européen formé par le CEPT (Commission Européenne des Postes et Télécommunication) et des industriels européens

- EUREKA est une structure européenne dont le but est d'encourager une plus grande collaboration entre industriels sur des projets de technologies avancées
- Eureka95 : projet TVHD de la Communauté Européenne qui implique Philips, Thomson, Thorn Emi et Bosch
- Eurotech : European Technical Advisory Committee to the Consumer Electronics Manufacturers - comité à vocation technique
- HDTV : High-Definition Television. A television system with approximately twice the horizontal and twice the vertical resolution of current 525-line and 625-line systems, component colour coding (e.g. RGB or YCbCr), a picture aspect ratio of 16:9 and a frame rate of at least 24 Hz
- IDTV : Improved Definition Television. A television system that offers picture quality substantially improved over conventional receivers, for signals originated in standard 525-line or 625-line format, by processing that involves the use of field store and/or frame store (memory) techniques at the receiver. One example is the use of field or frame memory to implement de-interlacing at the receiver in order to reduce inter-line twitter compared to that of an interlaced display. IDTV techniques are implemented entirely at the receiver and involve no change to picture origination equipment and no change to emission standards
- IEC : International Electrotechnic Committee
- JPEG : Joint Photographic Expert Group
- JTC1 : Joint Technical Committee 1, Information Technology (ISO/IEC)
- Letter-box : A television system that limits the recording or transmission of useful picture information to about three-quarters of the available vertical picture height of the distribution format (e.g. 525-line), in order to offer program material that has a wide picture aspect ratio
- NAB : National Association of Broadcasters
- NHK : Nippon Hoso Kyokai, principal radiodiffuseur japonais qui propose la norme tvhd japonnaise (voir MUSE)



- PICA : projet ESPRIT, Photographic Image Compression Algorithm
- RACE : Research for Advanced Communication in Europe
- QCIF :Quarter Common source Intermediatee Format (1/4 CIF, e.g. 1180\*144)
- Sample Aspect Ratio. The ratio of horizontal sample pitch to vertical sample pitch. A sample aspect ratio of unity achieves square pixels
- Square Pixels. A colloquial expression derived from computer graphics: a sample aspect ratio of precisely unity
- Self-contained ATV : A system that transmits a complete ATV signal in its own RF channel. A self-contained ATV system is channel-compatible if it transmits the ATV signal within a 6 MHz channel and receiver-compatible if it employs simulcasting. An example of a self-contained, non-channel-compatible system is NHK MUSE-9. Sometimes misleadingly called incompatible
- Side Panels : In a wide-screen augmentation system, those portions of the width of a wide aspect ratio picture that are not conveyed in the 4:3 aspect ratio
- UER : Union Européenne de Radiodiffusion
- WARC : World Administrative Radio Conference (1977->attribution bandes de fréquence et amplitude signaux TV pour canaux télévision)
- Wide-screen : A television system that offers a picture aspect ratio substantially wider than 4:3, using the full vertical picture area afforded by the distribution signal format (e.g. 525-line). Emission regulation changes are required for wide-screen terrestrial (UHF/VHF) TV transmission. Wide-screen may or may not be combined with EDTV; wide-screen is inherent in ATV and HDTV

## References

- [1] La TVHD américaine vers le tout numérique. *Electronique Internationale*, page 53, janvier 1991.

- 
- [2] M. Annegarne. HD-MAC : A Step Forward in the Evolution of Television Technology. *Philips Technical Review*, 43:pages 197–212, aout 1987.
  - [3] R. Besson. *Cours élémentaire de télévision moderne*. Edition Radio, 1976.
  - [4] CMTT. Transmission of Component-Coded Digital Television Signals for Contribution-Quality Applications at the Third Hierarchical Level of CCITT Recommendation G.702. 1990.
  - [5] J. Dumont. Les Techniques de la Télévision d'aujourd'hui à demain. *Dossiers de l'Audiovisuel*, pages 13–54, février 1988.
  - [6] F.W.P. Vreeswijk, W. Jonker, J.R.G.M. Leenen, J. van der Meer. An HD-MAC Coding System. In *2nd International Workshop on Signal Processing for HDTV*, L'Aquila, Italie, mars 1988.
  - [7] D. Le Gall. A Video Compression Standard for Multimedia Applications. *Communications of the ACM*, 34:pages 47–58, avril 1991.
  - [8] S. Matsumoto H. Murakami. Recent Trends in Bit-Rate Reduction Coding Methods for Broadcast TV and its International Standardisation. In *16th International T.V. Symposium*, pages 27–38, Montreux, Suisse, juin 1989.
  - [9] C.M. Carey-Smith Jr. F.W.P. Vreeswijk, M.R. Haghiri. HDMAC Coding for Compatible Broadcasting of High Definition Television Signals. In *16th International T.V. Symposium*, 1989.
  - [10] S. Okubo M. Yamashita, N. D. Kenyon. Standardization of Audiovisual Systems in CCITT. In *Image'Com 90*, pages 42–47, Bordeaux, novembre 1990.
  - [11] T.I.P. Trew O.J. Morris. Adaptive Spatial Sub-Sampling for HD-MAC. In *2nd International Workshop on Signal Processing for HDTV*, L'Aquila, Italie, mars 1988.
  - [12] Yves-Marie Le Pannerer. Various Aspects of HDTV Digital Transmission. In *Image'Com 90*, pages 66–71, Bordeaux, novembre 1990.
  - [13] T. Yamakawa S. Takagi. HDTV Telecommunication, Transmission System and Application. In *Image'Com 90*, pages pages 78–82, Bordeaux, novembre 1990.

- [14] G.K. Wallace. The JPEG Still Picture Compression Standard. *Communications of the ACM*, 34:pages 31–43, avril 1991.



---

Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,  
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY  
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex  
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1  
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex  
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

---

Éditeur  
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)  
ISSN 0249-6399